DOCUMENT RESUME

ED 073 426 CS 000 366

TITLE	Diagnostic Tests in Reading: An Annotated Bibliography.
INSTITUTION	New York State Education Dept., Albany. Bureau of Pupil Testing and Advisory Services.
PUB DATE	Mar 70
NOTE	61p.; Revised Edition
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Annotated Bibliographies; Reading; *Reading Diagnosis; *Reading Tests; *Test Reviews; *Test Selection

ABSTRACT

This publication lists and describes diagnostic reading tests available from United States publishers. Critical reviews by outside experts are included and supplemented by the compiler's comments. The external reviews are abstracted from "The Mental Measurements Yearbooks," edited by O. K. Buros, and cited by Yearbook edition and entry number. The introductory material serves two purposes: (1) to outline some considerations to be followed in selecting diagnostic tests and (2) to provide background for the understanding of technical information provided in test manuals and reviews. (Author/TO)

FILMED FROM BEST AVAILABLE COPY

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 073426

DIAGNOSTIC TESTS

IN READING

An Annotated Bibliography

The University of the State of New York
The State Education Department
Division of Educational Testing
Bureau of Pupil Testing and Advisory Services
Albany, N.Y. 12224

Revised Edition
March 1970

CS 000 866

Preface

This publication lists and describes diagnostic reading tests available from American test publishers. Critical reviews by outside experts have also been included and supplemented by the compiler's comments. The external reviews have been abstracted from The Mental Measurements Yearbooks, edited by O.K. Buros, and are cited by Yearbook edition and entry number so that further study of the reviews may be made. For example, MMY 6:829 refers to entry number 829 in The Sixth Mental Measurements Yearbook.

Attention is also called to the introductory material which is intended to serve two purposes: (1) to outline some considerations to be followed in selecting diagnostic tests and (2) to provide background for the understanding of technical information provided in test manuals and reviews.

It is hoped that all who use this bibliography will find it helpful in selecting tests which may provide useful clues to improving reading instruction.

This publication and three subsequent revisions (March 1968, September 1968, and March 1970) were prepared by Priscilla Hayward Crago of the Bureau of Pupil Testing and Advisory Services.

TABLE OF CONTENTS

Page

I. Preface i

II. Introduction. 1

How Do Survey Tests Differ from Diagnostic Tests?
What Types of Diagnostic Reading Tests Are Available?
Where Can You Find Additional Information?
How Do You Select Diagnostic Tests?
Meaningful Subscores
Subscore Reliability
Subscore Intercorrelations
What Other Factors Should Be Considered in Selecting Diagnostic Tests?
What about Norms?
Do Diagnostic Test Results Give the Whole Picture?

III. Diagnostic Tests (Word Attack and Phonics Components) 12

Adult Basic Reading Inventory
Basic Reading Inventory (cross-reference)
Botel Reading Inventory
California Phonics Survey
Classroom Reading Inventory
Delaware County Silent Reading Test, Second Edition
Denver Public Schools Reading Inventory
Diagnostic Reading Examination for Diagnosis of Special Difficulty in Reading
Diagnostic Reading Scales
Diagnostic Reading Tests
Doren Diagnostic Reading Test
Durrell Analysis of Reading Difficulty, New Edition
Durrell-Sullivan Reading Capacity and Achievement Tests
Gates-McKillop Reading Diagnostic Tests
Group Diagnostic Reading Aptitude and Achievement Tests
Inventory-Survey Test
Lee-Clark Reading Test, 1958 Revision
McCullough Word-Analysis Tests
Monroe Diagnostic Reading Tests (cross-reference)
Neale Analysis of Reading Ability
Phonics Knowledge Survey
Phonovisual Diagnostic Test
Primary Reading Profiles
Reading Skills Diagnostic Test
Roswell-Chall Auditory Blending Test
Roswell-Chall Diagnostic Reading Test of Word Analysis Skills
Silent Reading Diagnostic Tests: The Developmental Reading Tests
Standard Reading Inventory
Stanford Diagnostic Reading Test
Test of Individual Needs in Reading

IV. Work-Study Tests. 47
Appraisal of Reading Versatility
Developmental Reading Tests: Intermediate Level
Diagnostic Examination of Silent Reading Abilities
Iowa Silent Reading Tests, New Edition
Nelson Silent Reading Test
Pupil Progress Series Diagnostic Reading Tests
Reader Rater
The Reading Eye
Reading Versatility Test
SRA Reading Record

Introduction

Most teachers are familiar with survey reading tests which they regularly administer to their entire classroom. From such tests they learn where their class as a whole and where individual students within the class stand relative to some outside criterion. The outside criterion, a set of norms, is most commonly expressed as percentile ranks or grade-equivalents based on a national sample of pupils who took the same test. In some cases the survey test may provide several subscores, which give the teacher further insight into the class's strengths and weaknesses. For those pupils who score low on the survey test, another instrument is needed to pinpoint with greater precision the nature of the problems that are contributing to the overall low score so that remedial instruction can be appropriately planned and carried forth. The diagnostic reading test is designed to fill this need.

How Do Survey Tests Differ from Diagnostic Tests?

For the purpose of this bibliography, a distinction between survey and diagnostic tests was made on the basis of the number of subscores each test yielded. Survey tests may provide only a total score, or they may provide up to three subscores and a composite -- for a total of four scores. Typically the three survey subscores may be vocabulary (or word recognition), comprehension, and speed (or rate). Or the subscores may be word meaning (which corresponds to vocabulary) and sentence and paragraph meaning (which correspond to comprehension subscores on other tests). Some survey tests may tap these skills by requiring the pupil to read words and paragraphs orally, while others may require silent reading of the test content. Survey tests fill the need for ob-

taining broad assessment of entire classes and for classifying pupils as low, average, or high in reading achievement.

Diagnostic tests, on the other hand, are usually intended only for those pupils who have been screened first by a survey measure and identified as poor readers. By breaking down reading comprehension into its component skills and measuring each skill separately, the diagnostic test enables the teacher to concentrate remedial efforts on the skills that show relatively greater weakness in a particular child or group of children.

Diagnostic tests may yield four or more subscores. Subscores are not to be confused with item classifications provided by some survey tests. Such item classifications are based on the grouping of test questions which measure a common component. These groupings, when combined with other groups, yield a meaningful total score, but taken by themselves they are too small (frequency being based on only 3 to 10 items) to yield reliable subscores. Reliability is a function of test length, and therefore most diagnostic tests having real utility must necessarily be time-consuming.

What Types of Diagnostic Reading Tests Are Available?

Two major types of multi-score reading tests have been identified. In this bibliography the term "diagnostic" has been kept for those tests which yield scores on such reading components as word attack, auditory and visual discrimination, blending, vowel and consonant sounds, reversals, and other objectives usually found in reading instruction which emphasizes the acquisition of phonics skills.

The term "work-study" has been applied to a second category of tests. These tests yield scores on such components as skimming ability, use of

an index, and ability to read for main idea or for factual information. The work-study or critical reading tests emphasize the ability to adjust one's reading approach and speed according to the type of material (fiction, science, social studies, index, and dictionary), the length of the material (words, sentences, paragraphs, and extended passages), and the purpose (to get the main idea, retain details, and make inferences).

The sets of skill groups represented by these two types of tests may need to be diagnosed at different periods in the pupil's school career. The teacher in the primary grades may be concerned more with phonics skills which emphasize word recognition and word attack abilities, whereas the teacher of pupils who have already mastered word skills may wish to concentrate on developing critical reading skills or skills of the work-study type.

Where Can You Find Additional Information?

Although this bibliography may help to narrow down the choice, there are other sources of information to be consulted. The external reviews appearing here are not given in full, but may be found in The Mental Measurements Yearbooks* In reading reviews it should be borne in mind that those reviewers who are reading specialists may look primarily at the test content, whereas test experts may look at the technical characteristics of the test, such as information on reliability, validity and normative data.

Publishers' catalogs provide a further source of information. An up-to-date file of catalogs is a necessity to keep abreast of new test developments.

Finally, it is recommended that specimen sets be examined before making a decision. The test content and the manuals themselves provide

*Reading test entries from all yearbooks are now reproduced in one volume, Reading Tests and Reviews, edited by O.K. Buros, the Gryphon Press, 1968.

the best source of information of all. And an actual try-out of several tests may yield still further information.

How Do You Select Diagnostic Tests?

Three criteria are suggested as guidelines for reviewing diagnostic tests:

(1) How does the test measure the component skills, and do the subscores represent meaningful areas for providing remedial instruction?

(2) Are the subscore reliabilities sufficiently high (above .90) for individual use?

(3) Are the intercorrelations among subtests sufficiently low (below .65) to warrant differential diagnosis?

Since these criteria are technical ones, some explanation is in order.

Meaningful Subscores

The meaningfulness of subscores can only be determined by examining the test content directly. Some subscores, for instance, may be so global in nature that they do not provide much more information than the subscores available from a survey test.

In addition, several publishers may use the same subtest label and yet refer to entirely different abilities. For example, syllabication to some means identifying the precise point at which a word divides itself into syllables, whereas to others it may mean to count the number of syllables or to read aloud a polysyllabic word. Vocabulary may be measured by selecting a word that matches a picture, by selecting a word that fits a definition given orally or in writing, or by placing a word that best fits the context. Paragraph comprehension may be measured

with simple factual questions or with complex questions at or above the readability level of the paragraph on which the questions are based. It is therefore wise to look closely at the test and decide what skills the pupil is expected to demonstrate.

Furthermore, the teacher will want to determine which components should be diagnosed for children at different reading levels. For some the phonics skills will be of importance, and for others the work-study skills will be more meaningful in the particular situation.

Subscore Reliability

Reliability of a score indicates the consistency or stability of that score. No test has perfect reliability, since it is only a sample of all possible questions that could be asked. If the test contained a different sample of questions, different results might be obtained. The more questions included the more likely one can place confidence in the results. Thus the longer the test or subtest the more likely the reliability is to increase.

One statistic used for reporting the reliability of a test or subtest is the standard error of measurement. It is usually expressed in raw score units. The chances are two out of three that an obtained score differs from a true score within the score range defined by plus and minus one standard error. For example, a test with a standard error of 3 means that an obtained score of 25 probably lies between 22 (25 minus 3) and 28 (25 plus 3). Average scores for groups are subject to much smaller errors, since the errors in individual scores tend to compensate one another. For this reason, higher reliability and hence smaller standard errors are required when interpreting subscore patterns of individuals than when dealing with mean differences in a larger group.

Another reliability statistic that is more frequently reported is the correlation coefficient (r). It may range from -1.00 to +1.00 and may be computed in several ways. (1) Sometimes the test is split in half and the scores on each half correlated with each other to determine whether both halves rank the pupils consistently in the same way. The resulting split-half correlation is stepped up by the Spearman-Brown formula in order to estimate the reliability of the entire test. (2) Sometimes the results on one form of a test are correlated with the results on an alternate form to indicate that the two samples of questions are measuring in the same domain. (3) The test-retest method of estimating reliability determines whether the students' scores on the test are stable over time. (4) Still other methods, such as the Kuder-Richardson formulas, depend upon internal measures of consistency. Split-half and Kuder-Richardson coefficients, being based on internal measures from only one test administration, tend to be high and should not be applied to speeded tests. Alternate-form and test-retest coefficients tend to underestimate the actual reliability.

As a rule of thumb, a test is expected to yield reliability coefficients in the .80's to be satisfactory for group decisions and in the .90's for individual decisions. If the coefficient is lower, the scores might fluctuate by chance and an observed difference might not be a true difference at all. Diagnostic tests, because they are usually not used with an entire classroom but are for individualizing instruction for a single pupil, should meet the higher level of reliability in the .90's.

A correlation coefficient may be squared to indicate the percentage of common score variation that exists on the two sets of scores that

are correlated, i.e. the scores on two halves of a test, the scores on two alternate forms, etc. Thus a coefficient of .90 indicates consistent performance in 81 percent of the scores. The percentage of common variance or the consistency with which a test yields reliable scores decreases rapidly as the correlation coefficient decreases, as the table below shows.

Coefficient (r)	Percentage of Consistency (r^2)
.95	.90
.90	.81
.80	.64
.70	.49
.60	.36
.50	.25
.40	.16
.30	.09
.20	.04
.10	.01

Subscore Intercorrelations

It is generally considered that if scores on the component parts of a test are to be differentially useful, each should possess some unique significance and should be relatively independent of other components. The correlation coefficient, as it is used to estimate reliability, should be high to indicate that the halves or alternate forms or whatever are being correlated have much in common or are measuring in the same domain and are consistent over time.

When the correlation coefficient is used in studying the subscores, however, it should be low to indicate that the components are independent and unique. The higher the correlation between subscores the more likely it is that they are measuring the same thing. The lower the correlation between subscores the more likely it is that they are measuring different things.

When the subscore intercorrelation coefficient is squared, the percentage of interdependence or shared variance on the two subtests is obtained. On a reading test it may be expected that vocabulary and comprehension scores will be highly correlated. A hypothetical correlation of .70 indicates that almost half (49%) of the correlation may be accounted for by some type of overlapping in what is being measured. The remainder (51%) is unaccounted for and apparently unique to each subtest.

In diagnostic tests, if feasible, we want more than half of each subscore to be pure and independent of other subscores. If the intercorrelation between two subtests is .65, there is a 42 percent overlap (.65 squared) and 58 percent that does not overlap. A correlation coefficient of .65 or lower is considered a desirable criterion when differential treatment is indicated by subtests supposedly measuring independent skills.

In the case of diagnostic tests, which are intended for pupils of low achievement rather than for a larger population of widely varying ability, the desirability of low subscore intercorrelations may be of less importance than the other criteria of meaningful and reliable subscores. We may expect reading abilities of the student population as a whole to be intercorrelated above the .65 level, because good readers are generally good in all reading components and poor readers are generally poor. In diagnostic testing we may be less concerned with inter-individual differences and more with intra-individual differences. Within one individual one component skill may be mastered well and another not mastered at all. In the individual case, then, the correlation may be negative or low, even though the intercorrelations for a larger population were too high to meet the generally accepted criterion of .65.

What Other Factors Should Be Considered in Selecting Diagnostic Tests?

Other criteria for judging diagnostic tests are the same that apply to all tests, such as ease of administration and scoring, time, cost, and adequacy of interpretative information. Here, the ease of administration and scoring may depend upon the type of school staff that is available. If the school has a psychologist or a reading specialist with clinical training, a test that is difficult to administer may present no obstacles. If classroom teachers are expected to conduct the testing, the test chosen must necessarily be easier to administer and should yield subscores that lend themselves to direct interpretation in terms of concrete teaching objectives.

What about Norms?

Usually the norms for a test are in terms of percentile ranks or grade-level scores, based on a group of pupils similar to those for whom the test is intended. Ideally, the publishers of standardized tests will describe the norms sample carefully in terms of grade levels, geographic distribution, number of cases, socioeconomic status, community type, and general ability level. If norms are provided with a test, such information should be available. Unfortunately, these criteria are not universally met.

Some diagnostic tests do not provide norms on the premise that remedial efforts should be directed toward all pupils who show less than complete mastery of a component skill. Nevertheless, such tests do provide interpretive information. Instead of norms based on number right, they supply error analysis scores or checklists. The tests look at the number wrong rather than the number right. These tests then classify errors within meaningful categories. The goal becomes a perfect score or complete mastery rather than obtaining a relative rank.

This approach seems justified, although there should be evidence that the component scores are useful and that the test content has been pretested on a sample of students similar to the group for which it is intended.

As mentioned previously, the intra-individual pattern of scores is of prime importance in diagnostic testing. The differences between subscores for a single individual are inspected to find what areas should be stressed in remedial instruction. Since intra-individual profiles do not depend so heavily on outside criteria, we have another reason for accepting tests which lack norms. Here we may be less concerned with how the individual pupil deviates from the total population, for it is already known that his reading performance is low, and more concerned with how his own skills deviate among themselves.

Do Diagnostic Test Results Give the Whole Picture?

While diagnostic reading tests may provide clues to the content and level of remedial instruction, the competent reading specialist will take other factors into consideration. Classroom observation of oral and silent reading skills and habits will contribute to a better understanding of each child.

Information from other types of tests may also yield useful data. Specifically, tests of intelligence or scholastic ability, listening comprehension, and arithmetic computation are tools which can play an important part. If the scores on a survey reading test are on a par with those from a test in any one of these areas, the pupil may be reading at his optimum level for the time being. Remedial efforts may help him retain his relative standing. His progress will be slow as his reading component skills (perhaps also on a par with each other) improve at a steady pace.

If, on the other hand, the intelligence, listening, or arithmetic scores are considerably above a reading score, the child may be expected to benefit more quickly from a remedial program designed to correct his deficiencies. The score discrepancy often indicates that the pupil has potential to increase his reading skills to the level indicated by the higher scores on the other test or tests. The usual expectation is that a child who can deal with abstract figures and reasoning problems in an intelligence test, who can understand the spoken language in a listening test, or who can manipulate numerical symbols on an arithmetic computation test without word problems, can also handle letter symbols and words in their written form. This expectation is confirmed for the general school population by the high correlation that exists among intelligence, listening, reading, and arithmetic tests. Individual exceptions will, of course, occur. Some diagnostic reading tests include listening measures on the very premise that listening ability serves as a frame of reference to which reading ability can be compared.

Besides the use of teacher observation and testing techniques, other school resources may be used in an interdisciplinary approach. Home, school, and pupil factors may contribute to reading difficulties. The job of identifying motivation or emotional problems may lie with psychologists, parents, and others. The school librarian may have an important role in encouraging reading for enjoyment. Health factors such as hearing and vision problems may need to be investigated by the school nurse or a doctor. All in all, reading is a complex ability and requires the attention and dedication of many.

Title: Adult Basic Reading Inventory

Author: Richard W. Burnett

Publisher: Scholastic Testing Service, Inc., 480 Meyer Road, Bensenville, Ill. 60106

Date: 1966

Range: Functionally illiterate adolescents or adults (reading below the fourth grade level)

Forms: A

No. of Subscores: 6 subscores and total

Subscore Reliabilities above .90: Word Meaning (Reading), .97

Intercorrelations of Subscores: No information given.

Time: Under 60 minutes

Description: A group test which measures sight words, sound and letter discriminations (beginning consonants, blends, and digraphs), word meaning (reading), word meaning (listening), and context reading. If an individual gets chance scores on all but the listening test, he is classified as an absolute illiterate rather than a functional illiterate. The words in both word meaning subtests are identical and are of graded difficulty. When 70% accuracy is obtained when the words are read, the grade level indicates the instructional level. If the word meaning (listening) score is higher than the word meaning (reading) score, the individual is considered to have greater potential to benefit from instruction. Clues to basic skills strengths and weaknesses are found in the other subtests. Test booklets without the word "Adult" in the title are available for school use.

Norms: Standards rather than norms are provided. The instructional level standard is based on 70% accuracy on the word meaning (reading) score. The standard of complete mastery, i.e. a perfect score, is interpreted as indicating that the individual is not a functional illiterate.

External Review: None

Internal Review: Technical Report No. 1 provides useful validity and reliability information on the appropriateness of the word meaning (reading) score standard in placing adults at the proper instructional level. Additional technical reports are promised as test users share their research data with the publisher. The Manual contains valuable suggestions for score interpretation and the planning of reading instruction for literacy projects. Easy to administer and score.

Title: Basic Reading Inventory

SEE

Adult Basic Reading Inventory

Title: Botel Reading Inventory

Author: Morton Botel

Publisher: Follett Publishing Company, 1010 West Washington Boulevard, Chicago, Illinois 60607

Date: 1961-66

Range: Grades 1-12

Forms: A Battery and B Battery

No. of Scores: 4 ratings derived from 3 tests and 9 subscores obtained from fourth test.

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 39 to 107 minutes

Description: An individually administered test battery designed to classify reading level as Free Reading (pupil can read independently without teacher help), Instructional (pupil usually needs teacher guidance for comprehension), or Frustration (pupil cannot read with profit even with teacher help). These levels are determined by accuracy of responses to the Word Recognition Test, an oral reading test based on word lists of graded difficulty through the fourth grade level, and to the Word Opposites Test, which measures comprehension of written words of graded difficulty beyond the fourth grade level. The Word Opposites Test may also be administered as a group or individual listening test, in which case it yields a Potential Reading Level. The battery also includes a Phonics Mastery Test to evaluate knowledge of consonant sounds, consonant blends, consonant digraphs, rhyming words, long and short vowels, other vowel sounds, number of syllables, accented syllables and nonsense words. The nonsense words may be used to screen or double-check the phonetic elements covered in the other Phonics subtests.

Norms: Standards rather than norms are provided. The standard for the Phonics Mastery subtests is 100% accuracy. For the Word Recognition and Word Opposites tests 3 reading level standards are presented in grade-equivalent terms. For the Word Opposites (Listening) test a potential level standard is presented in grade-equivalent terms. Grade-equivalents range from pre-primer to senior high school. Grade-equivalents are not based on test results but on grade-placement of test content.

Botel Reading Inventory (cont'd)

External Reviews in MMY 6:834: Tests are appropriate only for pupils reading at grade 4 level or lower. Actual teacher experience has shown that the reading level standards are appropriate, even in the face of no supporting research evidence. One reviewer finds the test useful in spite of the lack of reliability and validity data, whereas the second reviewer is opposed to the use of the test because of the lack of technical data.

Internal Review: Attractive format. Use of Phonics Mastery subtests may yield useful clues for remedial instruction. A major drawback is the lack of statistical information. Users should be cautious in interpreting score differences, especially the difference between Instructional and Potential levels.

Title: California Phonics Survey

Authors: Grace M. Brown and Alice B. Cottrell

Publisher: CTB/ McGraw-Hill, Princeton Road S-2, Hightstown, N.J. 08520

Date: 1962-63

Range: Grades 7-12 and college

Forms: Form 1 (Form 2 is for survey purposes only and yields a total score).

No. of Scores: 8 error analysis subscores and total.

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 40-45 minutes

Description: A group test of the relationships between written and spoken words, which may be administered orally or by tape recording. Uses 5 listening and reading test exercises to reveal errors in 8 categories: Long-Short Vowel Confusion, Other Vowel Confusion, Consonant Confusion with Blends and Digraphs, Consonant-Vowel Reversals, Configuration, Endings, Negatives and Opposites, and Rigidity. Can be used below grade 7 if pupils have learned the mechanics of word recognition. A diagnostic key is used for scoring all 75 items, so that each item can measure more than one of the 8 diagnostic categories. Cutting scores are established on the total score to divide performance into 4 categories representing adequate phonics skill or one of 3 degrees of impairment. The authors state that "the definition of levels of phonic competence by means of the standard error of measurement...precludes the necessity for norms" and that, optimally, students who possess "adequate" phonics skill should make perfect scores.

Norms: Four categories of competence based on cutting scores applied to the Total score, but no norms for error analysis subscores. The standard errors of measurement used to locate the cutting scores are based on the normative sample consisting of students from 7 colleges, 16 elementary schools (public and private) in 11 states and in England. No information on size of sample.

External Review in MMY 6:820: Easy to administer. Adequate reliability and validity data resulted from research phase and standardization program. Carefully constructed. The use of the standard error of measurement to establish cut-off points is wise.

Internal Review: Reliability is adequate (.89-.93) for the Total score, but no data are reported for the 8 diagnostic scores. Manual contains good explanation of construction and validation of test. The use of nonsense words and of phonetic spellings in the test booklet makes the test free of criticism that it might be measuring conventional spelling.

Title: Classroom Reading Inventory

Author: Nicholas J. Silvaroli

Publisher: William C. Brown Book Co., 135 South Locust St., Dubuque, Iowa 52001

Date: 1965

Range: Grades 2-8

Forms: One form

No. of Scores: 6 subscores

Subscore Reliabilities above .90: Unknown

Intercorrelations of Subscores: Unknown

Time: Unknown

Description: Measures word recognition; independent, instructional, and frustration reading levels; hearing capacity level, and spelling.

Norms: Grade levels

Reviews: None. Information is incomplete, because a specimen set has not been examined.

Title: Delaware County Silent Reading Test, Second Edition

Authors: Judson E. Newburg and Nicholas A. Spennato

Publisher: Delaware County Reading Consultants Association, c/o Nicholas A. Spennato, Delaware County County Public Schools, Court House Annex, Media, Pa. 19063

Date: 1965

Range: Grades 1.5, 2.0, 2.5, 3.0, 3.5, 4,5,6,7,8

Form: One form

No. of Scores: 4 subscores and 1 total

Subscore Reliabilities above .90: Unknown

Intercorrelations of Subscores: Unknown

Time: Unknown

Description: Measures interpretation, organization, vocabulary, structural analysis and total

Norms:

Reviews: None. Above information is incomplete because a specimen set was not examined.

Title: Denver Public Schools Reading Inventory

Author: William R. Spears and Department of Instructional Services staff

Publisher: Department of General Curriculum Services, Denver Public Schools, 414 Fourteenth Street, Denver, Colorado 80202

Date: 1965-69

Range: Grades 1-8

Forms: One form

No. of Scores: 3 scores and ratings of 26 types of errors

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: Untimed

Description: An inidvidually administered test. Yields instructional, independent, and capacity reading level scores and rates a number of areas of strengths and weaknesses on the basis of oral responses to 12 vocabulary lists and 24 reading passages. The pupil reads the graded vocabulary lists until he makes 5 errors. The grade level for the passages to be read is determined by the level of the vocabulary list mastered successfully. The teacher records word recognition errors in 4 categories (omissions, insertions, substitutions, and punctuation) and answers to comprehension questions in 3 categories (factual, vocabulary meaning, and inference). There are also 12 word attack classifications (consonants, blends, etc.) and 7 areas of faulty approach (finger pointing, lip movement in silent reading, enunciation, etc.) to be observed.

Norms: Grade levels for major scores are presumably based on reading formulas and are computed from percentage mastery of word recognition (vocabulary) and comprehension (reading passage questions).

External Review: None

Internal Review: Lacks reliability data. Some subjectivity in scoring the comprehension questions is involved, with partial credit allowed for some answers. The rationale for the test appears sound, and the manual contains helpful suggestions for correcting deficiencies.

D

Title: Diagnostic Reading Examination for Diagnosis of Special Difficulty in Reading (or Monroe Diagnostic Reading Test)

Author: Marion Monroe

Publisher: C. H. Stoelting Company, 424 North Homan Avenue, Chicago, Ill. 60624

Date: 1928-29

Range: Grades 1-4

No. of Scores: 14 subscores and 12 error analysis categories

Subscore Reliabilities above .90: No information given.

Intercorrelation of Subscores: No information given.

Time: About 2 hours

Description: An individual test, requiring practice and special training to administer. Nine analytic tests yield 14 subscores. Test 1, Alphabet Repeating and Reading, requires the pupil to say the alphabet and to identify letters on stimulus cards. It is scored for time to say the alphabet, number of letters omitted, and number of letters incorrectly identified on the cards. Test 2, Iota Word Test, requires the child to read orally the words on the stimulus cards to determine grade level. Test 3, b, d, p, q, u, n Test, uses a stimulus card which the child reads orally; it is scored for interchanging the letters u and n or the letters b, d, p, and q. In Test 4, Recognition of Orientation, the child is required to distinguish orally between forward and reversed letters and words. Test 5, Mirror Reading, requires the use of a book and mirror; the child reads from the mirror image. This test is scored for the time to finish the passage, the number of words misread, the ratio of time for mirror reading and normal reading, and the ratio of errors made in the two ways of reading. Test 6, Mirror Writing, requires the child to mirror write dictated words and is scored for the number of letters written correctly. Test 7, Number Reversal, requires the pupil to read numbers from a stimulus card; it is scored for reversal and orientation errors. In Test 8, Word Discrimination, the child points to words on the stimulus cards which match the key word; it is scored for number correct. In Test 9, Sounding, the examiner says the sounds of words slowly and the child orally tells what the words are; it is scored for number correct.

Errors on Tests 2 and 8 (Iota Word Test and Word Discrimination) and on Gray's Oral Paragraphs may be analyzed and summarized in a profile which groups errors in 14 categories (faulty vowels, faulty consonants, reversals, etc.) to show whether the child makes errors typical of his grade placement or has excessive errors.

Diagnostic Reading Examination for Diagnosis of Special Difficulty in Reading (or Monroe Diagnostic Reading Test) (cont'd)

An educational profile may also be made, using grade - or age - equivalents from mental ability, arithmetic, spelling, and other oral and silent reading tests by the author. This profile show whether the child is retarded in reading only or in or spelling, too, with scores adjusted for chronological and Information on handedness may also be gathered.

Norms: Interquartile range, 75th and 90th percentile for grades 1-3 for Tests 3, 4 and 5. Grade-equivalents for Tests 2, 8, and 9. "Excessive error" ratings for Test 7. Interquartile range for error analysis categories. No description of norming group.

External Review: None

Internal Review: No reliability or validity information is provided. The norms group is not described in terms of size or other characteristics. Some of the information in the mimeographed manual, such as references to other tests and publishers, is out of date. The manual rightfully stresses importance of training on the part of the examiner and the need to establish rapport in testing a child with a reading disability. Although not mentioned in the manual, it is evident that the examiner should possess knowledge of the emotional and central nervous system disorders connected with some reading disabilities. The test appears carefully constructed to serve the need of identifying children with special reading disabilities Even though more complete and up-to-date norms are needed, the test may be useful in the hands of a clinician.

D

Title: Diagnostic Reading Scales

Author: George D. Spache

Publisher: CTB/ McGraw-Hill, Princeton Road S-2, Hightstown, N.J. 08520

Date: 1963

Range: Grades 1-8 and retarded readers in grades 9-12.

Forms: One form

No. of Scores: 11 subscores

Subscore Reliabilities above .90: Word Recognition (.87-.96). Other subscore reliabilities are either lower or not reported at all.

Intercorrelations of Subscores: No information given.

Time: 45 minutes.

Description: An individually administered test requiring the pupil to read aloud or silently, while teacher records the pupil's responses. Initial tests (Word Recognition, Oral Reading, and Silent Reading followed by comprehension questions) yield grade equivalents, which serve as a basis for determining level of achievement at which supplementary phonics tests (Consonant Sounds, Vowel Sounds, Consonant Blends, Common Syllables, Blending, and Letter Sounds) are to be administered. A test of Auditory Comprehension is also included. From the Oral Reading, Silent Reading, and Auditory Comprehension tests, respectively, are derived scores of Instructional Level, Independent Level, and Potential Level. A Rate of Silent Reading score may also be obtained. The supplementary phonics tests measure abilities to relate printed symbols to sounds.

Norms: Grade-equivalents (ranging from 1.3 to 6.5) for Word Recognition, Instructional Level, Independent Level, and Potential Level. Ratings of slow, average, and fast for Rate score. Average scores for 6 supplementary phonics tests are reported for 1, 2 and 3 years of phonics training. No description of norming group.

External Reviews in MMY 6:821: Easy to administer and score. Auditory comprehension is not rightfully a measure of "potential reading level" inasmuch as norms show that listening comprehension is 2 grade levels above grade placement on the average. Four major scores are useful when accompanied by clinical analysis of errors.

Internal Review: The norms groups which determined the grade level placement in the initial tests are not described. The lack of norms (other than average scores) for the supplementary phonics tests indicates that perfection is the goal; for these supplementary tests, no reliability data are reported.

Title: Diagnostic Reading Tests

Author: Committee on Diagnostic Reading, Frances Oralind Triggs, Chairman

Publisher: Committee on Diagnostic Reading Tests, Inc., Mountain Home, N.C. 28758

Date: 1947-1966

Range: K-4 level; Lower Level: Grades 4-8; Upper Level: Grades 7-13

Forms: K-4 level: A and B; Lower Level: A, B, C, and D; Upper Level: A, B, C, and D for the Diagnostic Battery and A, B, C, D, E, F, G, and H for the Survey Section. For some subtests there are fewer forms than indicated.

No. of Scores: K-1: 10 subscores and 3 subtotals; Grades 2-4: 2 subscores and 1 subtotal; Lower Level: 21 subscores and 3 subtotals; Upper Level: 29 subscores and 7 subtotals.

Subscore Reliabilities above .90: Eye-Hand and Motor Coordination (.93, K-1); Visual Discrimination (.96, K-1); Vocabulary (.90-.95, K-1 and 4-8); Auditory Discrimination (.93, Grade 1); Word Attack Survey (.91, Grade 2); Word Recognition and Comprehension subtotal (.90-.93, Grades 2-6); Silent Word Attack: Sounds and Syllabication subscores and subtotal (.93-.97, Grades 4-13); Social Studies Vocabulary (.94, Grades 7-13); Survey Total (.91, Grades 7-13). Information not given for all subscores.

Intercorrelations of Subscores: No information given.

Time: 40 to about 80 minutes for the Survey Sections; untimed Diagnostic Sections (except Rate of Reading).

Description: At the K-4 level the Reading Readiness and Survey Sections for Kindergarten and Grade 1 measure Relationships, Eye-Hand and Motor Coordination, Visual and Auditory Discrimination, and Vocabulary, with several subscores within each category. Survey Booklet II for Grade 2 and Survey Booklet III for Grades 3 and 4 measure Word Recognition and Comprehension. Booklet IV, consisting of paragraphs and word lists of graded difficulty, is an individually administered oral Word Attack test to spot errors in 7 categories: Substitutions, Omissions, Repetitions, Mispronounciations, Insertions, Pauses, and Refusals.

The Survey Section at the Upper Level covers Comprehension, Vocabulary, and Rate of Reading, whereas the Lower Level covers the same areas plus Word Recognition. At the Upper level weaknesses in any of these areas may then be assessed by Diagnostic Sections covering the same areas in greater detail. At both levels a Word Attack test is available for diagnosing oral and silent skills in matching Sounds and Syllabication. The Vocabulary section provides scores on Words per Minute; General Vocabulary; Vocabulary Comprehension; and specialized English Literature, Mathematics, Science, and Social Studies Vocabulary. The Rates of Reading section yields separate measures of Speeded and Unspeeded Rate of General Reading, Social Studies, and Science materials. The Comprehension section may be given as a silent or auditory test.

Diagnostic Reading Tests (cont'd)

Norms: Quartiles and medians for the Readiness Test subscores, based on 35 children in a nonpublic kindergarten, tested in the spring; 932 children in a suburban grade 1 school, tested in Fall 1957; and 962 children in a southern semi-rural community, tested in Fall 1957. Percentiles for other tests. Numbers of pupils on which norms are based vary from one subscore to another, with a range from 188 to 16,600. The norms are based on scores submitted voluntarily to the Committee and are revised when the addition of new data changes the form of the curve in a direction more representative of public schools. Some norms tables were tabulated in 1951, 1953, 1956, and 1959 and others bear revision dates of 1959 and 1963.

External Reviews in MMY 6:823 and MMY 4:531: One reviewer considers the tests generally useful, while another advocates a "good overhaul of the tests and norms. Lower and Upper Levels are more satisfactory than the K-4 level. Manuals should be consolidated. Some items in the Diagnostic Section are identical to those in the Survey Section, and some items are faulty. Reliability data are incomplete and too low for individual diagnosis. Rates of Reading section yields separate rate scores for science and social studies which is unwarranted since the test does not reliably measure any flexibility in rate.

Internal Review: The grade level of the children in the semi-rural Readiness norms group is not indicated. The composition of the norms groups at the higher levels is not described. The Word Attack test, which might yield the most important information for remedial phonics work, classifies errors in less meaningful categories than those found in other tests. At the higher grade levels the emphasis is on study skills, but the coverage is incomplete. The test may be more useful for survey than for diagnostic purposes, since the subtotal scores are sufficielty reliable and the content validity is satisfactory.

D

Title: Doren Diagnostic Reading Test of Word Recognition Skills

Author: Margaret Doren

Publisher: American Guidance Service, Inc., Publishers' Building, Circle Pines, Minnesota 55014

Date: 1956

Range: Grades 1-4 or above

Forms: One form

No. of Scores: 11 subscores and total

Subscore Reliabilities above .90: What the author reports as reliability is actually validity, or the correlations of the parts to the whole (see below).

Intercorrelations of Subscores: Part score correlations with total test range from .53 to .88. No information given on subtest intercorrelations.

Time: About 3 hours.

Description: A group test measuring the following word recognition skills: Letter Recognition, Beginning Sounds, Whole Word Recognition, Words Within Words, Speech, Consonants, Ending Sounds, Blending, Rhyming, Vowels, Sight Words, and Discriminate Guessing. Each skill is subdivided into finer classifications. Success on test depends on use of visual, auditory, and context clues.

Norms: No norms for subtest scores, but mean total scores are reported for grades 1-4.

External Reviews in MMY 5:659: Content is based on word recognition skills emphasized in widely used basic reading series. Some faulty items. Author prefers analysis of errors rather than norms to indicate level of achievement. Norms group not described.

Internal Review: Lacks reliability data. The directions to the pupil may be difficult to understand, but examples are used to clarify. Author omits norms (based on number right), saying that number wrong and the type of errors made are the important factors for diagnostic testing. Manual is clearly written and gives concrete suggestions for remedial activities. Profile for recording errors makes use of "line of significant error" (7 errors per subtest) below which a need for remediation is indicated; no information is given on its statistical derivation. Test gives comprehensive coverage of word recognition skills and is recommended in spite of the limitations noted.

D

Title: Durrell Analysis of Reading Difficulty, New Edition

Author: Donald D. Durrell

Publisher: Harcourt, Brace & World, Inc., 757 Third Avenue, New York, N.Y. 10017

Date: 1955

Range: Grades 1-6

Forms: One form

No. of Scores: 8 subtests and checklist of errors

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 30 to 90 minutes

Description: Individually administered test, consisting of two sets of graded paragraphs, one for oral and one for silent reading, which are scored on the basis of time required to complete. Level of comprehension is measured with factual questions. Additional subtests measure Word Recognition and Word Analysis, Letters (for non-reader), Visual Memory of Words or Word Forms (for non-readers or those reading below grade 3 level), Auditory Analysis of Word Elements, Spelling and Handwriting, and Listening Comprehension. Word Recognition is administered with a cardboard tachistoscope. Diagnosis depends almost entirely on checklists of difficulties. Checklists for medical, psychological, and educational records are also included.

Norms: Time scores (oral and silent reading) are rated as Low, Middle, and High at each grade level, 1 to 6. Number of silent reading memories elicited by the examiner are also rated as L, M, and H. Author states that the norms are "based on no fewer than 1,000 children for each test" and that the checklists of errors are more important than the norms.

External Reviews in MMY 5:660: No tests of word attack methods, e.g. syllabication or phonic analysis, are included. Clear instructions. Paragraphs well chosen and graded. Lack of norms for paragraph reading comprehension is a drawback. Tests are inadequate for sampling reading below the third grade level. The norms sample is not described. No reliability or validity data. Listening comprehension is included to measure potential for reading comprehension in non-readers. Checklists may be most useful part of the tests in the hands of reading teachers and clinicians.

Internal Review: The testing situation is not standardized, because the examiner is given leeway in the questions he asks to measure factual comprehension and in the timing of Word Recognition which uses a tachistoscope.

Title: Durrell-Sullivan Reading Capacity and Achievement Tests

Authors: Donald D. Durrell and Helen Blair Sullivan

Publisher: Harcourt, Brace & World, Inc., 757 Third Avenue, New York, N.Y. 10017

Date: 1937-45

Range: Primary: Grades 2.5-4.5; Intermediate: Grades 3-6

Forms: A and B (Intermediate Achievement part only); A only for Intermediate Capacity test and for both Primary parts.

No. of Scores: 6 subscores and 2 total scores.

Subscore Reliabilities above .90: In the Capacity Test, Word Meaning (.84-.91) subscore reaches acceptable level. In Achievement Test both Word Meaning (.90-.96) and Paragraph Meaning (.83-.93) subscore reliabilities are satisfactory. No information given on optional Spelling and Written Recall subtests.

Intercorrelation of Subtests: No information given.

Time: 73-90 minutes

Description: Reading Capacity Test (a measure of hearing comprehension) yields scores on Word Meaning, Paragraph Meaning, and Total, and is contrasted to the Reading Achievement Test, which yields similar scores. Optional tests of Spelling and Written Recall are provided.

Norms: Age-equivalents and grade-equivalents for totals and all subscores except Spelling. Grade-equivalents only are available for Spelling. The Intermediate Form A norms were based on 6,000 pupils in 19 communities in 11 states, in grades 3-6. Form B norms were based on equating with Form A. No information is given on the derivation of Primary norms.

External Reviews in MMY 5:661 and 4:562: Based on assumption that responses to hearing comprehension test and pictorial material and to a reading achievement test of equivalent difficulty can detect reading disability if there is a discrepancy. Validity of this premise is not established. Manual confuses objective data with sales approach. Earlier review was favorable, considering the test a useful way to identify poor readers who are capable of reading beyond present level.

Internal Review: May identify poor reading but not diagnose the types of difficulties. The same criticism may hold true for the "Durrell Listening-Reading Series", copyrighted 1968, and available from the same publisher. The series is not listed here, since it yields fewer subscores--2 vocabulary scores (listening and reading) and two sentence comprehension scores (listening and reading). The technical information, however, is more complete for the new series, including normative and intercorrelation data.

Title: Gates-McKillop Reading Diagnostic Tests

Authors: Arthur I. Gates and Anne S. McKillop

Publisher: Teachers College Press, Teachers College, Columbia University, New York, New York 10027

Date: 1962

Range: Grades 2-6

Forms: I and II

No. of Scores: 28 subscores

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: About 60 minutes, assuming all parts are seldom given.

Description: An individually administered test. The examiner records the pupil's oral responses in most parts, but the pupil marks his own answers to other parts. The examiner is given leeway as to whether to administer all parts or not. The first section requires oral reading of paragraphs of graded difficulty, from which a grade-equivalent score is derived; errors may be classified in 11 ways (omissions, additions, repetitions, mispronunciations, reversals, wrong beginning, wrong middle, wrong ending, etc.). In the next three parts the child pronounces words and phrases of increasing difficulty, presented with a hand cardboard tachistoscope and under untimed conditions so that the teacher can note methods of attack. The fifth part (Knowledge of Word Parts: Word Attack) proceeds from large units (nonsense words) to small units (letter sounds and naming capital and lower case letters). If the child makes errors when looking at a whole nonsense word, the examiner shows parts of the words to see if the child can blend them into wholes. In the sixth section (Recognizing the Visual Forms or Word-Equivalents of Sounds) four subtests require the student not to supply sounds as in the previous word attack tests but to recognize sounds heard and to associate them with their visual forms by marking multiple-choice items; the examiner proceeds from nonsense words to requesting the child to select the initial letters, final letters, and vowels in the middle of words spoken. In the seventh part the examiner says parts of words slowly and the child tells what the word is in order to measure auditory blending skill. If the test is too difficult for the child, the examiner is instructed to administer the Auditory Discrimination test. The eighth section consists of four supplementary tests of oral spelling, oral vocabulary, syllabication, and auditory discrimination. In the syllabication subtest the child reads nonsense words of two or more syllables, and in auditory discrimination he is required to indicate whether pairs of words spoken by the examiner are the same or different.

Gates-McKillop Reading Diagnostic Tests (cont'd)

Norms: Grade-equivalents for Oral Reading scores (paragraphs, words, and phrases) and supplementary spelling and vocabulary scores. Other raw scores or error scores are rated as high, medium, low, or very low or as normal progress, low, and very low in terms of either the pupil's actual grade placement or his oral reading grade-equivalent. No norms for auditory discrimination scores. No information on norming sample.

External Reviews in MMY 6:824: The tests include no measure of reading comprehension. After the oral reading part, future subtest scores are interpreted with respect to the oral reading level, thus leading to continual bias and under-or over-estimating the level of difficulty in subsequent scores. Requires sophistication to interpret.

Internal Review: No reliability or validity data. The norms group is not described in terms of size, grade range, or any other characteristics. The test content, however, reflects careful construction. The use of nonsense words to measure word attack skills is to be commended. Administration and scoring procedures seem very difficult and require practice to ensure mastery of the recommended techniques. For instance, the examiner needs to know which tests are to be completed and which are to be discontinued after a specified number of errors. The stimulus material sometimes appears in the teacher's booklet and sometimes in the pupil's. The examiner is also instructed to use certain symbols for scoring errors as the test proceeds and to be observant of guessing behavior in the multiple-choice sections. Scoring sometimes depends on correct responses and sometimes on errors, and one part calls for prorating of raw scores if the entire subtest is not given.

D

Title: Group Diagnostic Reading Aptitude and Achievement Tests

Authors: Marion Monroe and Eva Edith Sherman

Publisher: C. H. Nevins Printing Company, 811 Bryn Mawr Island, Bayshore Gardens, Bradenton, Florida 33505

Date: 1939 (re-copyrighted 1966)

Range: Grades 3-9

Forms: One form

No. of Scores: 15 subscores

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: About 50 minutes

Description: A group test. The Achievement section consists of subtests measuring Paragraph Understanding, Speed, Word Discrimination (Vowels, Consonants, Reversals, Additions and Omissions), Arithmetic and Spelling. The Aptitude section has subtests of Visual, Auditory, and Motor Ability and Vocabulary. Visual Memory tests require the pupil to reproduce nonsense words and forms shown on a card exposed for 5 or 10 seconds. The Auditory tests require the pupil to reproduce letters given orally and to identify the position of the word in a list of four words. In the Motor tests the pupil copies a written paragraph and crosses out a's in nonsense words. In Vocabulary the child underlines one of four phrases presented orally and in writing which makes sense.

Norms: Grade-equivalents for some of the subtests in the Achievement section and percentile ranks for all other subtests. No description of the normative population. Norms are considered tentative.

External Review: None

Internal Review: In Paragraph Understanding the questions precede each paragraph, an approach which seems good for measuring directed reading or scanning. The Speed subtest requires the pupil to write short answers as he reads, and thus the speed is cut down and seems a questionable practice. The other subtests appear valid and attractive. The mimeographed set of norms includes suggestions for remedial work. The possible causes of errors and suggested remedies are clearly explained in concrete terms, contributing much to the value of the test in a diagnostic, remedial setting. In spite of the tentative nature of the norms, the old publication date, and the lack of reliability data, the test appears useful for the purpose it is intended to serve.

Title: Inventory-Survey Test

Author: Marion Monroe

Publisher: Scott, Foresman & Company, 99 Bauer Drive, Oakland, New Jersey 07436

Date: 1965

Range: Grades 4-6

Forms: A (B in preparation)

No. of Scores: 5 subscores and total

Subscore Reliabilities above .90:

Intercorrelations of Subscores:

Time:

Description: A group test developed as part of the Scott, Foresman Reading Program, it consists of 225 items measuring Word Meaning, Sentence Meaning, Total Meaning, Word Analysis, and Dictionary Skills. Total Meaning uses paragraphs and extended stories followed by comprehension questions. Word Analysis measures a number of phonics components such as consonant and vowel sounds, syllabication, and root words but does not yield separate scores for the components. In Dictionary Skills the components measured include using a pronounciation key, krowledge of tense and plurals, alphabetizing, and again no component subscores are available.

Norms: Grade equivalents ranging from 1.5 to 9.0.

External Review: None

Internal Review: The test manual was not available for review, and thus some of the information above is incomplete. The material within each subtest is arranged in order of difficulty, with many of the items in Word Meaning appearing particularly difficult below the high school and college level. However, a raw score of 15 out of 25 is given a grade-equivalent of 5.0. Since no breakdown of Word Analysis skills is provided, the teacher will want to follow-up low scores with a more diagnostic test or make an item analysis of responses in this potentially useful subtest.

Title: Lee-Clark Reading Test, 1958 Revision

Authors: J. Murray Lee and Willis W. Clark

Publisher: CTB/ McGraw-Hill, Princeton Road S-2, Hightstown, N.J. 08520

Date: 1958

Range: Primer level: Grade 1; First Reader level: Grades 1-2

Forms: A and B

No. of Scores: Primer: 3 subscores and total; First Reader: 5 subscores and total.

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 22 minutes

Description: Primer and First Reader levels both have subtests on Auditory Stimuli (the ability to hear, remember and relate spoken sound to symbol), Visual Stimuli (word recognition or the ability to match words to pictures), and Following Directions (sentence comprehension or the ability to mark pictures in a prescribed way). Additional subtests of the First Reader include Completion and Inference, which require the pupil to select a word that best completes a sentence. Considered by authors to be a survey test rather than a diagnostic measure. Test booklets are available in i/t/a.

Norms: Grade-equivalents, age-equivalents, and percentiles for Total score. Means for subscores. Norms provided separately for a representative national sample (Group I) and for an above average sample (Group II) which began formal reading instruction early in grade 1. Group I norms were obtained by "comparison" with California Reading Test results. Group II norms were obtained by "comparison" with results from the 1943 edition of the Lee-Clark Primer and First Reader tests "and other tests".

External Reviews in MMY 6:795: A revision of the 1943 Edition. Helpful discussions of test interpretation appear in the manuals. Users should treat subscore differences with caution because reliability data and intercorrelations of parts are not provided. Easy to administer and score. One reviewer criticizes provision of "average" and "above-average" norms groups as being inadequately defined and difficult to apply. One favorable and one unfavorable review recommending that tests not be used.

Internal Review: Manual gives evidence of validity and comparability of alternate forms but ignores reliability. Norms group is not adequately described, and norming procedures are not clearly explained. Does "comparison" mean that an equating program took place, or were the tests re-normed?

Title: McCullough Word-Analysis Tests

Author: Constance M. McCullough

Publisher: Ginn & Company, Statler Building, Park Square, Boston, Massachusetts 02117 (Also available from Personnel Press, 20 Nassau Street, Princeton, New Jersey 08540)

Date: 1962-63

Range: Grades 4-6

Forms: One form

No. of Scores: 7 subscores, 2 subtotals, and total.

Subscore Reliabilities above .90: Initial Blends and Digraphs (.96-.97), Phonetic Analysis subtotal (.93-.94), and Total (.94-.96). All other subscore reliabilities are low.

Intercorrelations of Subscores: No information given.

Time: 70 minutes

Description: A group test. First 5 subtests (Initial Blends and Digraphs, Phonetic Discrimination, Matching Letters to Vowel Sounds, Sounding Whole Words, and Interpreting Phonetic Symbols) yield separate subscores and a combined Phonetic Analysis subtotal. The last 2 subtests (Dividing Words into Syllables and Root Words in Affixed Forms) yield separate subscores and a combined Structural-Analysis subtotal. The last 3 subtests are intended for pupils reading at the fifth grade level and beyond.

Norms: Percentiles for boys, girls, and total. About 600 pupils at each grade, 4-6, from 23 school systems (identified by name) in 21 states, participated in the November-December 1962 norming program. They were selected to be representative of socioeconomic status distribution in the general population and had an average IQ of 108.

External Reviews in MMY 6:826: Scores cluster near the tops of the distribution; for example, a perfect raw score of Test 1 has a percentile rank in the 60's. Test 3 requires knowledge of proper spelling (rather than variant spellings of a sound) of word read aloud by the teacher. Two reviewers agree that the test as a whole is well-constructed and useful.

Internal Review: The manual frankly points out the especially low reliability of the Dividing Words into Syllables subtest and the clustering of scores at the top of some distributions. Author suggests remedial work for pupils who score one standard deviation below the mean on a given subtest. Subtests appear to have good content validity, but their reliabilities are low.

Title: Monroe Diagnostic Reading Test

SEE

Diagnostic Reading Examination for Diagnosis of Special Difficulty in Reading

Title: Neale Analysis of Reading Ability

Author: Marie D. Neale

Publisher: St. Martin's Press, Inc., 175 Fifth Avenue, New York, New York 10010

Date: 1958-66

Range: Ages 6-12

Forms: A, B, and C

No. of Scores: 3 major scores, 3 supplementary scores, 6 error analysis categories, and a checklist of several qualitative factors.

Subscore Reliabilities above .90: Word Accuracy (.96-.98) and Comprehension (.92-.98)

Intercorrelations of Subscores: No information given.

Time: 10 to 15 minutes

Description: An individually administered test or oral reading, developed for use in Britain. From six prose passages of graded difficulty followed by questions, scores on comprehension, accuracy (lack of reading errors), and word-per-minute rate are obtained. Errors may be further classified as mispronunciations, substitutions, refusals, additions, omissions, and reversals. The alternate forms are spiral-bound in one booklet. Three supplementary diagnostic tests deal with the names and sounds of letters, auditory discrimination through spelling, and blending. A checklist is included to aid in observing attitudes, reading habits, and word recognition approaches.

Norms: Age-equivalents for the accuracy, comprehension, and rate scores. The norms are based on children aged 7 to 11 from 13 schools in Britain. Form A was administered to 1221 children, Form B to 552, and Form C to 489. The sample was controlled for age, sex, social background, and urban-rural type of community. No norms for the supplementary tests.

External Reviews in MMY 6:843: There are no directions for administering, scoring, or interpreting the three supplementary tests. The information on construction and standardization is scanty. The sixth passage is too difficult. The norms beyond age 11 are extrapolated, and percentile rank norms would be helpful. The best feature of the test is the categorization of errors on standard reading passages.

Neale Analysis of Reading Ability (cont'd)

Internal Review: No reliability data are reported for the rate and supplementary test scores. The test booklet is attractive. The test appears easy to administer. Some British spellings (centre and programme in Form A and honours and favour in Form C) are present. The qualitative checklist and and analysis of errors and supplementary test scores may provide more useful diagnostic information than the part which offers normative data and which differs little from a survey test.

D

Title: Phonics Knowledge Survey

Authors: Dolores Durkin and Leonard Meshover

Publisher: Teachers College Press, Teachers College, Columbia University, 525 West 120th Street, New York, New York 10027

Date: 1964

Range: Grades 1-6

Forms: One form

No. of Scores: 15 subscores

Subscore Reliabilities above .90: No information given.

Intercorrelation of Subtests: No information given.

Time: 10-30 minutes

Description: Individually administered test. Teacher poses questions orally with use of Content Card and marks pupil's errors (unknown sounds, wrong pronunciations, etc.) on Response Record. Designed to determine what students know about specific elements of phonics in order to plan instruction accordingly. Yields 15 scores on: Names of Letters, Consonant Sounds, Vowels: Long and Short, Vowel Generalizations, Sounds of C and G, Sounds of Y, Consonant Blends, Digraphs, Vowel Combinations, Vowels Followed by R, Sounds of QU, Sounds of OO, Sounds of X, Beginning Consonant Combinations, and Syllabication. No norms. No reliability data.

Norms: None

External Review: None

Internal Review: Subtests appear too short for sufficient reliability. Lack of norms implies that perfect score is goal for each child to achieve and that any errors detected indicate area in which remedial instruction is to be given.

Title: Phonovisual Diagnostic Test

Authors: Lucille D. Schoolfield and Josephine B. Timberlake

Publisher: Phonovisual Products, Inc., P.O. Box 5625, Friendship Station, Washington, D.C. 20016

Date: 1949-58

Range: Grades 3-12

Forms: 1949 edition and 1958 edition

No. of Scores: Analysis of more than 26 error categories

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 15 minutes

Description: Measures phonetic weaknesses by analysis of spelling errors in list of 17 or 20 words dictated by teacher. Error analysis can be made of beginning and final consonant sounds, digraphs, 9 initial blends, and 17 vowel sounds.

Norms: None

External Reviews in MMY 6:829: Originally a spelling test, combining a measure of auditory discrimination and knowledge of spelling conventions. (Cannot be sure which of 2 factors is being measured in each child.) Spelling could be learned by children lacking phonic skills, and test is therefore of doubtful validity. Alternate forms are not equivalent. The average grade-placement of words is 4.0 to 4.9, thus making the test too easy for some.

Internal Review: Technically a weak test, lacking validity, reliability, and norms data. However, it seems to cover the most improtant elements of phonics. The lack of norms assumes that a perfect score is the goal and that the types of errors indicate the areas in which remedial instruction is to be given. Could be useful, quick device for use by the classroom teacher, but a reading specialist would want to follow up with a more sophisticated test.

Title: Primary Reading Profiles

Authors: James B. Stroud, Albert N. Hieronymus, and Paul McKee

Publisher: Houghton Mifflin Company, 53 West 43rd Street, New York, N.Y. 10036

Date: 1953-68

Range: Level 1: End of grade 1; Level 2: End of grade 2

Forms: One form

No. of Scores: 5 subscores and 1 composite

Subscore Reliabilities above .90: Auditory Association (.92-.96), Word Recognition (.91-.95), and Word Attack (.86-.92) at both levels. Reading Comprehension (.95-.96) at Level 1 but not at Level 2 (.84-.87).

Intercorrelations of Subscores: Word Attack and Reading Comprehension correlate too highly (.82) to be differentiated from each other at Level 2. Word Recognition and Word Attack correlate too highly at Level 1. Other subscores more likely to have diagnostic value.

Time: 87 minutes for Level 1; 79 minutes for Level 2.

Description: The score on the Aptitude for Reading subtest, a measure of listening, is to be compared with the Reading Comprehension score. If the scores diverge indicating actual reading ability is lower than expected reading ability, the remaining scores (Auditory Association, Word Recognition, and Word Attack) are inspected for possible clues of difficulty. In Auditory Association, the pupil matches word sounds to pictures. In Word Recognition, he matches a printed word to the word read aloud by the teacher. In Word Attack, the pupil chooses a word omitted in a paragraph, using context and auditory clues.

Norms: Percentile ranks. Level 1 norms are based on 1,067 pupils from 53 schools in 21 states; Level 2 norms are based on 1,056 pupils from 28 schools in 18 states, tested in May, 1956.

External Reviews in MMY 5:665: Aptitude for Reading subtest, says one reviewer, is not valid as an aptitude measure because its correlations with the rest of the battery and with a readiness test are low, and its reliability is also low (.61-.77). Other reviewer likes the inclusion of the Aptitude subtest and praises the entire battery. Tests 2-5 have good content validity.

Internal Review: Reliabiltiies of most subscores are encouraging, although the intercorrelation of Word Attack with other skills is discouraging. Perhaps the best differentiation of skills can be made between Auditory Association and the two word scores taken together (Recognition and Attack). This test may, therefore, be used diagnostically to indicate whether difficulties lie in the auditory or written domain. Additional diagnosis may be needed to point up specific phonic disabilities.

D

Title: Reading Skills Diagnostic Test

Author: Richard H. Bloomer

Publisher: Brador Publications, Inc., Livonia, New York 14487

Date: 1967

Range: Grades 2-8

Forms: One form

No. of scores: 9 subscores

Subscore Reliabilities above .90: Phonetic Words (.91) and Inconsistent Words (.92)

Intercorrelations of Subscores: No information given.

Time: Untimed

Description: An individually administered test, using a test booklet and set of flash cards, it measures letter identification, letter-sound identification, phonetic sounds, phonetically consistent words, phonetically inconsistent words, consistent phrases, inconsistent phrases, letters in context, and words in context. Each successive skill tested is assumed to be dependent upon mastery of the preceding skills. As a result, the test is not administered in entirety and remediation of errors is recommended before administering later parts of the test. Some responses are made orally and some in writing.

Norms: Reading grade level norms for Letters in Context (grades 2.0 to 6.0) and for Words in Context (grades 4 to 8). No norms for other subtests.

External Review: None

Internal Review: The subtest reliabilities are spuriously high because they are based on a total grades 1 to 6 sample instead of on a sample with a more restricted range. The manual includes specific lessons and games for teaching linguistic remedial reading. The lack of norms except for the final two parts implies that perfection is the goal.

Title: Roswell-Chall Auditory Blending Test

Authors: Florence G. Roswell and Jeanne S. Chall

Publisher: Essay Press, P.O. Box 5, Planetarium Station, New York, N.Y. 10024

Date: 1963

Range: Grades 1-4

Forms: one form

No. of Scores: 1 total score

Subscore Reliabilities above .90: Total score, .86-.94

Time: About 5 minutes

Description: An individual, orally administered test designed to measure the ability to blend sounds into whole words, with the purpose of judging whether the pupil will have ease or difficulty in later phonics instruction. The teacher presents the sounds of each word slowly, and the child responds orally by telling what the word is.

Norms: Ratings of "inferior" or "adequate" are assigned to raw scores for pupils in grades 1-5. The cutting score at each grade level was determined by testing 62 New York City Negroes in grade 1, retested again in grades 2-4, and 25 severely disabled readers in grades 3-5.

External Reviews in MMY 6:830: The dichotomy of rating appears inadequate and is based on nonrepresentative samples of children. Interpretation should be in terms of prediction of success in phonics instruction, since this is the stated purpose. The test has almost none of the characteristics of good test construction and validation.

Internal Review: The usefulness of a test which yields so little information for planning instruction seems doubtful.

Title: Roswell-Chall Diagnostic Reading Test of Word Analysis Skills

Authors: Florence G. Roswell and Jeanne S. Chall

Publisher: Essay Press, P.O. Box 5, Planetarium Station, New York, N.Y. 10024

Date: 1956-59

Range: Grades 2-6

Forms: I and II

No. of Scores: 6 subscores

Subscore Reliabilities above .90: Short Vowel Sounds (.99), Vowel Combinations (.93).

Intercorrelations of Subscores: No information given.

Time: About 5 minutes

Description: An individual test to which the pupil responds orally. He is required to sound single consonants, consonant combinations, short vowels given in isolation and in one syllable words, words with silent e, vowel combinations, and words of more than one syllable. The last subscore is called syllabication, although the pupil does not need to isolate each syllable.

Norms: None. A "few errors within a subtest" indicate need for instruction. If "more than half the items on any subtest" are missed, the child has a "special deficiency in this area" or "has not received instruction" and needs it.

External Reviews in MMY 6:831 and MMY 5:667: A child may pronounce words correctly because of a large sight vocabulary or word attack skills other than knowledge of sounds. The only way to check functional use of word attack skills is to confront the child with unknown or nonsense words. Reliability and validity data are not in the manual but are available from the publisher on request. The test is of limited use as a measure of word recognition skills because interpretative data are missing. Test does not give adequate coverage of some common blends and fails to measure ability to apply knowledge of isolated vowel sounds in unfamiliar contexts. Some sounds are tested in words likely to be in the child's sight vocabulary. Test appears to measure certain phonetic elements rather than skills in word analysis as the title implies.

Internal Review: Easy to administer, but the interpretation of scores is left up to the teacher, with only subjective guidelines offered by the publisher. Since the reliability data reported are based on a heterogeneous group (ranging from grade 1 to 8), the coefficients are inflated. Reliability within a single grade would be much lower.

Title: Silent Reading Diagnostic Tests: The Developmental Reading Tests

Authors: Guy L. Bond, Theodore Clymer, and Cyril J. Hoyt

Publisher: Lyons & Carnahan, 407 East 25th Street, Chicago, Ill. 60616

Date: 1955

Range: Grades 3-5 or 3-8

Forms: D-A (One form)

No. of Scores: 15 subscores and 5 composite subscores

Subscore Reliabilities above .90: Recognition Pattern No. Right composite score (.97), Visual Analysis composite score (.94), Locating Elements (.95), Phonetic Knowledge composite score (.95).

Intercorrelations of Subscores: No information given.

Time: 90 minutes

Description: A group test of silent reading skills, consisting of 11 subtests. From the Word Recognition section (Recognition of Words in Isolation and in Context, Recognition of Reversible Words in Context) 6 Recognition Pattern scores (Total Right, Total Errors, Total Omitted, Words in Isolation, Words in Context, and Orientation) and 4 Error Analysis scores (Initial, Middle, Ending, and Orientation Errors) are derived. From the Recognition Techniques section 4 Visual Analysis scores (Locating Element, Syllabication, Locating Root Word, and Visual Analysis composite) and 5 Phonetic Knowledge scores (Word Elements, Beginning Sounds, Rhyming Sounds, Letter Sounds, and Phonetic Knowledge composite) are obtained. The remaining subtest, Word Synthesis, yields the 20th score. Scoring keys are provided for classifying the type of error (beginning, middle, etc.) which occurs in the first two Word Recognition Tests. The Developmental Reading Tests series also includes survey tests, suitable for grades 1-6, which emphasize comprehension.

Norms: Grade-equivalents ranging from 1.5 to 6.0, described as tentative, were based on results of "all the children in a typical Midwestern community with a population of approximately 20,000 and ... other groups from (2) other representative communities."

External Reviews in MMY 6:832: Test 4 and 6 items could have several correct answers. One reviewer finds scoring keys awkward to use, while another reviewer finds them easy to use. Good to excellent for appraising word recognition and word attack skills, the tests do not measure comprehension. Subtest reliabilities high enough for individual diagnosis.

Internal Review: Subtest reliabilities are <u>not</u> high enough for individual diagnosis, with 4 exceptions. Tentative norms group is not adequately described in terms of grade levels, size, and other

Silent Reading Diagnostic Tests: The Developmental Reading Tests (cont'd)

factors. With reliability data based on fifth graders for one subtest and on third graders for all other subtests, one wonders whether only the more satisfactory data have been presented. The reliability of subscore differences is discussed in terms of "Lines of Importance" which mark off a central portion on the profile to introduce some caution into score interpretation. The Manual and profile suggest remedial instruction only if reading results are "appreciably below" mental grade, thus taking each individual's intelligence level into account.

D

Title: Standard Reading Inventory

Author: Robert A. McCracken

Publisher: Pioneer Printing Company, 306-B Flora, Bellingham, Wash. 98225

Date: 1966

Range: Pre-primer through seventh reader levels

Forms: A and B

No. of Scores: 4 reading levels and 8 achievement area scores

Subscore Reliabilities above .90: Reliabilities range from .68 to .99, with only 4 of the 12 scores below a reliability of .90.

Intercorrelations of Subscores No information given.

Time: Untimed

Description: An individually administered test, measuring independent, maximum instructional, minimum instructional, and frustration reading levels. Each form has 11 stories of graded difficulty for oral reading, 8 stories for silent reading and 11 word lists (presented with a hand tachistoscope) for measuring word pronouncing ability in isolation. An independent reading level is established by use of the word lists. Then the child begins with an oral story at his estimated level and proceeds to more difficult oral and silent stories until his frustration level is reached. Oral reading and word recognition errors, time, and comprehension and silent reading time and comprehension are evaluated along the way. There is an optional listening comprehension test for measuring potential level, and a checklist of reading habits.

Norms: Reading grade level standards for independent, instructional and frustration levels. The vocabulary, word recognition errors, speed, and comprehension achievement areas are similarly rated.

External Review: None

Internal Review: The high subtest reliabilities reported are misleading in view of the fact that they are based on a total grades 1 to 6 sample instead of a more restricted range. Scoring of oral word recognition errors seems complicated. Comprehension questions are not completely standardized. Since the main purpose is to place children within their appropriate instructional level and to measure speed and comprehension, the test appears more like a survey test than a diagnostic test. Word recognition errors are rated in a global way rather than broken down into diagnostic categories.

Title: Stanford Diagnostic Reading Test

Authors: Bjorn Karlsen, Richard Madden, and Eric F. Gardner

Publisher: Harcourt, Brace & World, Inc., 757 Third Avenue, New York, N.Y. 10017

Date: 1966

Range: Level I: Grades 2.5-4.5; Level II: Grades 4.5-8.5

Forms: W and X

No. of scores: 7 subscores for Level I; 7 subscores and 1 composite subscore for Level II.

Subscore Reliabilities above .90: Reading Comprehension at both levels (.87-.95), Auditory Discrimination at Level I (.94-.96), Beginning and Ending Sounds (.87-.92) at Level I, Blending at both levels (.94-.97), Sound Discrimination at both levels (.91-.94).

Intercorrelations of Subscores: Blending and Beginning and Ending Sounds subtests of Level I are too highly correlated (in .80's) to have diagnostic value. Vocabulary and Reading Comprehension are too highly correlated in Level II. Other subscores are moderately correlated and therefore independent of each other, affording diagnostic usefulness. Rate of Reading in Level II is the most independent measure.

Time: 137 minutes for Level I and 91 minutes for Level II.

Description: A group test with subtests of Reading Comprehension, Vocabulary, Syllabication, Sound Discrimination, and Blending at both levels. Additional subtests at Level I include Auditory Discrimination and Beginning and Ending Sounds. Level II provides a breakdown of Reading Comprehension into Literal and Inferential Comprehension and includes a Rate of Reading subtest. Vocabulary, Auditory Discrimination, and Level I Blending are administered orally. Pupils are required to relate sounds and pictures to written words and syllables, correctly spelled.

Norms: Stanines, scaled scores, and percentile ranks. Norms for Form X are based on samples of 100 which were drawn randomly from each grade from a total tested population of 12,000 pupils, the entire enrollment in six school systems, grades 2-8, tested in October 1965. The sample was drawn to be representative of all reading ability levels, as measured by the Stanford Achievement Test: Reading. Form W norms were derived by equi-percentile equating.

External Review: None.

Stanford Diagnostic Reading Test (cont'd)

Internal Review: Carefully constructed and standardized. Vocabulary, Syllabication, and Rate of Reading subscores lack sufficient reliability for individual diagnosis. Blending may be measuring spelling rather than blending of spoken sounds (Level I) or blending of written sounds (Level II); for example, "mb" is keyed as the correct ending sound for the word "lamb". Authors state that the subtests are more reliable at the lower levels of performance where reliability is most needed. Reading Comprehension, a sentence completion excercise, is similar to other tests which measure effectiveness of expression. The subtests which are most reliable and free of faults are Auditory Discrimination and Sound Discrimination.

FILMED FROM BEST AVAILABLE COPY

Title: Test of Individual Needs in Reading, Sixth Edition

Author: Hap Gilliland

Publisher: Montana Reading Clinic Publications, Eastern Montana College
517 Rimrock Road, Billings, Montana 59102

Date: 1961-66

Range: Grades 1-7

Form: One form

No. of Scores: 19 [illegible] total

Subscore Reliabilit[illegible] formative given

Intercorrelation[illegible] formative given

Time: About 40-50 [illegible] silent reading, 5-10 for oral reading, about 15-20 f[illegible]

Description: The st[illegible]et is entitled "John Bidwell and the Trail to [illegible]sisting of graded passages (primer to grade 7 and[illegible]tell a connected story, interspersed with comprehe[illegible] to be individually administered as an oral test [illegible]red as a silent reading test. The independent re[illegible]letermined when the student reaches a passage where he makes [illegible] errors per 100 words orally read. This section yields up to 3 scores (oral or silent reading level, comprehension, and rate). A separate word analysis section, administered individually with oral responses, yields 16 subscores and a subtotal (use of context, words beginning alike, beginning consonants, ending consonants, consonant substitution, speech consonants, consonant blends, reversals, long and short vowels, vowel blends, blending letter sounds, prefixes, suffixes, compound words, recognizing syllables, syllabication, and word analysis subtotal).

Norms: Grade levels based on reading formulas, texts, and sample of 2500 children attending the reading clinic (700 at remedial levels), and equating to two standardized tests.

External Review: None

Internal Review: No reliability information is given. The directions for administering and scoring are somewhat difficult to understand.

Title: Appraisal of Reading Versatility

Authors: Arthur S. McDonald, Sister Mary Alodia, George Zimny, Stanford E. Taylor, and James Byrne

Publisher: Educational Developmental Laboratories, Inc., 284 Pulaski Road, Huntington, L. I., New York 11744

Date: 1960-62

Range: Basic: Grades 6-10; Advanced: Grades 11-16 and adults

Forms: Basic: AA and BB; Advanced: CC and DD

No. of Scores: Basic. 7 subscores; Advanced: 6 subscores

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 15-30 minutes.

Description: Non-secure forms of the Paper and Pencil Editions of the Reading Versatility Test (q.v.), available to non-professionals and designed for self-administration. Measures reading flexibility by comparing a pupil's rates of reading under different sets of instructions. Refer to description of the Reading Versatility Test for further information.

Norms: See Reading Versatility Test

External Review: None

Internal Review: See Reading Versatility Test

FILMED FROM BEST AVAILABLE COPY

Title: Developmental Reading Tests: Intermediate Level

Authors: Guy L. Bond, Theodore Clymer, and Cyril J. Hoyt

Publisher: Lyons & Carnahan, 407 East 25th St., Chicago, Ill. 60616

Date: 1955-65

Range: Grades 4-6

Forms: A and B

No. of Scores: 5 sub[illegible] score

Subscore Reliabilities [illegible]own

Intercorrelations of Sc[illegible]own

Time: 32 minutes

Description: A group [illegible]ary levels for grades 1 to 3 measure vocabulary and [illegible]comprehension and may be considered survey tests. T[illegible] level measures basic vocabulary, reading to retai[illegible] reading to organize, reading to evaluate-inter[illegible] appreciate, and average comprehension. Thus, th[illegible]level appears to measure work-study skills.

Norms: Grade-equivalents

External Reviews in MMY6:787: Two reviewers give mostly negative comments for all levels of the tests. At the time there was no teacher's manual or information on norming, reliability, or validity. (There now is a manual for the Lower and Upper Primary Tests, supplying some of the missing information). The Intermediate level subtests are too short for satisfactory reliability, except the Reading to Organize part. The Intermediate forms are not of equivalent difficulty, with Form A easier than B. Chance scores may convert to grade-equivalents too high. Attractive format, easy to administer and score.

Internal Review: The manual for the Intermediate level still seems to remain unavailable. Consequently there is no way to judge the technical characteristics of the test.

WS

Title: Diagnostic Examination of Silent Reading Abilities

Authors: August Dvorak & M. J. Van Wagenen

Publisher: Van Wagenen Psycho-Educational Laboratories, 1729 Irving Avenue South, Minneapolis, Minn. 55411

Date: 1939-54

Range: Grades 4-6, 7-9, 10-13

Form: One form

No. of Scores: 11 subscores and 1 subtotal

Subscore Reliabilities above .90: None at highest level

Intercorrelations of subscores: .75 to above .90

Time: About 105-160 minutes

Description: The test consists of 4 parts: (1) Part 1, Rate of Comprehension; (2) Part 2, measuring 4 areas (perception of relations, vocabulary in context, vocabulary in isolation, and information). (3) Part 3, measuring 5 areas (central thought, single details, related ideas, inferences, and interpretation) and a subtotal; (4) Reading for Ideas, yielding a rate of reading score.

Norms:

External Reviews in MMY3: 480: There are signs of careless construction and keying of some items (more than one correct choice). An outside study of intercorrelations, corrected for attentuation, reported a range from .75 to above .90, indicating lack of uniqueness for more than half the subscores. Useful as measure of reading level and speed, but of little diagnostic use because of low reliabilities and high intercorrelations. Easy to score.

Internal Review: Above information is incomplete, because a specimen set has not been examined.

FILMED FROM BEST AVAILABLE COPY

Title: Iowa Silent Reading Tests, New Edition

Authors: Harry A. Greene, A. N. Jorgensen, and Victor H. Kelley

Publisher: Harcourt, Brace & World, Inc., 757 Third Avenue, New York, N.Y. 10017

Date: 1943

Range: Elementary: Grades [illegible]; Advanced: Grades 9-13

Forms: AM (Revised), BM [illegible] and DM

No. of Scores: 8 subscores and [illegible] for Elementary; 9 subscores and total for Advanced.

Subscore Reliabilities [illegible] grade 4, Alphabetizing in grades 6, 7 and 8. [illegible] subscores reaches a satisfactory level.

Intercorrelations of [illegible] Alphabetizing has lowest correlation with other Elementary [illegible] is therefore the most independent measure; however, it contributes little to the total test score. Most of remaining intercorrelations at the Elementary level are sufficiently low. At the Intermediate level, all intercorrelations are low, with the Rate score the most independent of other measures.

Time: 49 minutes for Elementary; 45 for Intermediate.

Description: A group test designed to measure silent reading proficiency of the work-study type. The Elementary and Advanced [illegible] yield subscores on Rate, Comprehension, Directed Reading, Word Meaning, Paragraph Comprehension, and Sentence Meaning. At the Elementary Level additional scores include Location of Information: Alphabetizing and Use of Index. At the Advanced level additional scores include Poetry Comprehension, Use of Index, and Selection of Key Words.

Norms: Standard scores, percentiles, grade-equivalents, and age-equivalents, based on pupils from areas "widely distributed geographically" tested in 1942. No indication of size or other characteristics of the norming group.

External Review in MMY 6:794: There has been no revision of the test since 1943 except for supplementary directions for separate answer sheets. Some of the content is outdated. The norms group is old and not adequately described. Not recommended.

Internal Review: Reading Rate and Alphabetizing are the only scores having satisfactory reliability. More scores should be reliable to warrant the use of the test. The outdatedness of the norms and some items also makes the test of minimum usefulness as a work-study test.

Title: Nelson Silent Reading Test

Author: M. J. Nelson

Publisher: Houghton Mifflin Company, 53 West 43rd Street, New York, N.Y. 10036

Date: 1931

Range: Grades 3-9

Forms: A, B, and C

No. of Scores: 4 subscores and total

Subscore Reliabilities above .9): No information given on subscores. Total score reliability (.87-.93) reaches acceptable level.

Intercorrelations of Subscores: No information given

Time: 30 minutes

Description: A group test consisting of vocabulary and paragraph comprehension items. From the paragraph section, three subscores are obtained: General Significance, Note Details, and Predict Outcome.

Norms: Age-equivalents and grade-equivalents, based on results from 41,000 pupils "from various parts of the United States". No further description of the norming group is provided.

External Reviews in MMY 4:545: Outdated manual. No statistical validation that test can be used diagnostically.

Internal Review: Grade 3 and 4 median scores are chance scores, indicating that the test is too difficult at these grade levels. Since no information is given on subscore reliability and the norms are so old, the test is also not recommended for grades 5-9.

FILMED FROM BEST AVAILABLE COPY

Title: Pupil Progress Series Diagnostic Reading Tests

Authors: Oliver F. Anderhalter, R. Stephen Gawkoski, and Ruth Colestock

Publisher: Scholastic Testing Service, Inc., 480 Meyer Road, Bensenville, Ill. 60106

Date: 1956-65

Range: Primary I: end of grade 1 - early grade 2; Primary II: grades 2-3; Elementary: grades 4-6; Advanced: grades 7-8

Forms: A and B

No. of Scores: 6 to 10 subscores and 3 subtotals

Subscore Reliabilities above .90: Vocabulary subtotal (.92-.97, grades 2-3); Rate subtotal (.93-.95, grades 1-2 and 4-8); Reading Comprehension subtotal (.75-.96, grades 1-7); Knowledge and Use of Sources subtotal (.82-.93, grades 7-8); Words in Use (.84-.95, grades 2-3); Word Meaning (.81-.94, grades 2-6 & 8); Reading to Locate Information (.74-.92, grades 4 & 6).

Intercorrelations of Subscores: Vocabulary and Comprehension subtotals intercorrelate in .70's and .80's, indicating considerable overlap. Most subscores are independent to a satisfactory degree.

Time: 40 minutes

Description: A group test. All levels yield Reading Comprehension and Rate subtotals. The Primary I and II Levels both measure Vocabulary, whereas the Elementary and Advanced Levels both measure Knowledge and Use of Sources (with subscores in Functions of Common Sources, Selection of Suitable Sources, Use of Index, and Use of Table of Contents). Vocabulary subscores include Word Recognition, Word to Content Relation, Words in Use, and/or Word Meaning. Comprehension subscores, depending on the level, include Recalling Information, Locating Information, Reading for Directions, and/or Word Meaning. The tests are designed to measure general reading ability in 3 major areas and to supplement this information with scores from the diagnostic subtests which emphasize ability to locate rapidly material of different types and to retain and understand what is read.

Norms: Percentiles for subscores and subtotals. Grade-equivalents for subtotals only. Form A norms based on sample of 11,510 pupils coming from 37 schools in 8 Western, Midwestern and Eastern states, tested in May, 1956, selected to be representative of all types of school systems. Form B norms based on 12,199 pupils tested in May, 1957. Adequacy of norms rechecked in 1960 and 1964.

External Review in MMY6:822: See page 53

External Review: (cont'd)

The reliability of part scores is questioned. For a diagnostic test, it should include auditory or word attack skills (the reason the internal reviewer classifies it as a measure of work-study or critical reading skills). The comprehension questions give undue weight to factual details at the primary and elementary levels.

Internal Review: The 1965 Technical Report gives clear explanation of construction, reliability, validity, and norming data, and honestly points out cautions to be observed in test interpretation. The reliabilities for the speeded tests (Rate, Primary I and Advanced Locating Information, 5 tests in the Elementary level, and 3 in the Advanced level) are inflated. None of the Knowledge and Use of Sources subscores and very few of the other subscores are reliable enough for differential diagnosis of individual pupils. Form B has more ceiling than Form A. Alternate forms show considerable differences in their reliability. Since the test is so short and the subscores so unreliable, it would be better to consider this test for survey purposes than for pinpointing the nature of strengths and weaknesses.

FILMED FROM BEST AVAILABLE COPY

Title: Reader Rater

Author and Publisher: Better Reading Program, Inc., 230 East Ohio Street, Chicago, Ill. 60611

Date: 1965

Range: Age 15 and over

Forms: Book 1 and Book 2

No. of Scores: 12 subscores and total

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 60-120 minutes

Description: A self-administered survey, developed as part of the Rapid Reading Kit. Books 1 and 2 are intended, respectively, as pre- and post-tests for the program. Yields scores on Speed, Comprehension, Reading Habits, Reading for Main Idea, Reading for Details, Reading for Related Ideas, Reading for Inference, Summarizing, Skimming, Recall, Unspeeded Vocabulary, Speeded Vocabulary, and Total. Answers are pin-punched in a Self-Scoring Profile. Comprehension and Recall depend on multiple-choice questions relating to a passage. Speed is in terms of words per minute. Speed rating and scores for other subtests are changed to percentages, which in turn are rated in 6 categories from poor to superior.

Norms: 6 qualitative categories ranging from poor to superior. No description of norming population from which such categories were derived.

External Review: None

Internal Review: Norms sample is not described. The same percentage scores convert to the same rating in Book 2 (the post-test) as in Book 1 (the pre-test), although it might be wise to have "stiffer" norms for Book 2, based on students who have pursued the program. The Raters are available separately and would be useful in classes that have not pursued the program. The self-administering feature makes the test attractive for older pupils. Teachers may find the profiled results helpful for planning instruction in work-study reading skills.

Title: The Reading Eye

Authors: Stanford E. Taylor and Helen Frackenpohl

Publisher: Educational Developmental Laboratories, Inc., 284 Pulaski Road, Huntington, L. I., New York 11744

Date: 1959-60

Range: 8 levels: Grades 1, 2, 3, 4, 5, 6, 7-8, 9-16 and adults

Forms: 8 forms at each of 8 levels

No. of Scores: 10 subscores

Subscore Reliabilities above .90: Reading component scores reported to have reliability between .83 and .93. No information given for other scores.

Intercorrelations of Subscores: No information given.

Time: 4 minutes

Description: Complete kit including eye-movement camera and test materials costs $700. Individually administered test. Pupil reads several test selections of graded difficulty aloud until a level which can be read with reasonable fluency is found. A passage at this level is then read silently and followed with 10 questions. With an eye-movement camera, a photographic record is obtained of oculomotor efficiency in the reading act. The record is analyzed to yield 5 Reading Component scores (fixations, regressions, average span of recognition, average duration of fixation, and rate with comprehension), 3 Ratings (grade level of fundamental reading skill, relative efficiency, and directional attack), and 2 Diagnostic Categories (visual adjustment and general adjustment to reading).

Norms: Grade-level norms for reading component scores only, given separately for males and females, based on a nationwide sample of 12,000 pupils, chosen to be representative of public and nonpublic schools in all geographic areas.

External Reviews in MMY 6:838: Alternate forms and normative data make retesting and study of progress from grade 1 through college practicable. Validity of the eye-movement camera technique has been demonstrated by correlations in the .80's and .90's between photographic measurements and scores on standard reading tests. Directions for use of camera and score interpretation are clearly written for teacher use. Two reviews concur with high praise for the material and technique.

Internal Review: None. Manual and test materials not inspected.

Title: Reading Versatility Test

Authors: Arthur S. McDonald, Sister Mary Alodia, George Zimny, Harold M. Nason, Stanford E. Taylor, and James Byrne

Publisher: Educational Developmental Laboratories, Inc., 284 Pulaski Road, Huntington, L. I., N. Y. 11744

Date: 1960 68

Range: Basic: Grades 5-8 or 6-10, Intermediate: Grades 8-12, Advanced: Grades 11-16 and adults.

Forms: Basic: A and B; Advanced: C and D

No. of Scores: Paper and Pencil editions: 7 subscores and 1 total; Basic Reading Eye Edition: 29 subscores

Subscore Reliabilities above .90: No information given.

Intercorrelations of Subscores: No information given.

Time: 15 to 50 minutes

Description: In the Basic Paper and Pencil Editions, the test measures reading flexibility by comparing a pupil's rates of reading under different sets of instructions. In Part 1 he is to read carefully in order to answer questions posed on a later page. In Part 2 he is to read a fictional story fast to understand its main points in order to answer questions posed later. In Part 3 he is to scan to find an answer to a question given in advance. In Parts 4 and 5 he is to skim the previous passages from beginning to end to answer questions given in advance which concern the main ideas of the passage. In the Advanced level, skimming is reduced to one part. These tests yield 2 Rate of Reading, 2 Comprehension, 1 or 2 Skimming Rate, and 1 Scanning Rate scores. Skimming and scanning rates are compared to regular rate of reading by means of a ratio to determine whether the pupil exhibits flexibility.

The same test passages may be used with the eye-movement camera in the Reading Eye Edition (Basic level only). The eye-movement photographs are analyzed to provide scores on fixations, regressions, span of recognition, duration of fixations, and rate within the five categories: careful reading, rapid reading, scanning, and 2 skimming exercises. Additional scores are Relative Efficiency and Directional Attack (left to right).

Norms: Grade level averages (grades 1-12 and college) given for Raw score in Paper and Pencil Editions. Also given for all Reading Eye Edition subscores. Publisher states that standardization is currently under way to provide grade equivalents and percentiles. Information is found in EDL Bulletin No. 3 (1960) rather than in the manual.

Reading Versatility Test (cont'd)

External Review: None.

Internal Review: Lacks reliability data and well-defined norms. Alternate forms said to be equated, but no evidence is presented. Clever approach for measuring critical reading skill of flexibility. Attractive format, but answer blocks for some comprehension questions on left-hand page appearing on right-hand page may be difficult to handle.

FILMED FROM BEST AVAILABLE COPY

Title: SRA Reading Record

Author: Guy T. Buswell

Publisher: Science Research Associates, Inc., 259 East Erie Street, Chicago, Ill. 60611

Date: 1947-59

Range: Grades 6-12

Forms: one form

No. of Scores: 4 subscores and total derived from 10 subtests

Subscore Reliabilities above .90: Everyday Reading Skills (.82-.93). Other subscores do not reach satisfactory level.

Intercorrelations of Subscores: Reading Rate score is most independent measure. Vocabulary score is most highly correlated with other scores, indicating that vocabulary contaminates other scores and that they cannot be differentiated apart from another. Factor correlation of the 10 subtests was the method used for grouping the tests to yield 4 subscores.

Time: 26 minutes

Description: Subtest 1 (Rate of Reading, a count of words read in two minutes) yields Reading Rate score; Subtest 2 (Reading Comprehension based on questions relating to article in Subtest 1) yields Comprehension score; Subtests 4-7 (paragraph, directory, map-table-graph, advertisement, and index reading) yield Everyday Reading Skills score; Subtests 8-10 (technical vocabulary, sentence meaning, and general vocabulary) yield Vocabulary score.

Norms: Percentiles for subscores and grade-equivalents for Total only. Based on equi-percentile equating with other SRA Reading Tests results in grades 6, 9, and 12, tested in the spring of 1957. Norms in intermediate grades were obtained by linear interpolation. Equating roups (310 to 331 pupils per grade) were chosen for being representative of national school population in terms of sex, age, socioeconomic status, and intelligence.

External Review in MMY 4:550: Short, highly speeded test, easy to administer. Adequately measures vocabulary, but broader skills such as interpretation and inference are largely neglected. Other comments in review are not pertinent since new norming and the grouping of 10 subtests to yield 4 subscore instead of 3 as formerly.

Internal Review: Of doubtful use because of low reliability of all subscores except one.